U-NET CONVOLUTIONAL NEURAL NETWORK FOR TEM IMAGE SEGMENTATION

Štefan Mocko

Master Degree Programme (2), FEEC BUT

E-mail: xmocko00@stud.feec.vutbr.cz

Supervised by: Tomáš Potočňák

E-mail: xpotoc00@stud.feec.vutbr.cz

Abstract: This work deals with the use of a convolutional neural network in the area of segmentation of images acquired with the use of a transmission electron microscope. Paper describes programming tool for image data augmentation, used neural network topology, and it also provides information about model training. This neural network topology delivered excellent results on provided data from the Thermo Fisher Scientific company, which will serve as a starting point for internal company research in image segmentation area.

Keywords: Neural network, U-net, CNN, TEM, segmentation

1 INTRODUCTION

Variability of TEM samples requires highly sophisticated algorithms. One possible solution to cope with this variability is the use of convolutional neural networks [4], [1]. One of the properties of neural networks is an ability to learn. Neural network training for semantic segmentation [5] of samples with a specific property, manufacturing defect, contamination, etc. could significantly reduce the microscope operator's search time.

2 METHODS

2.1 DATA AUGMENTATION

Original image data were acquired on Thermo scientific TEM systems, 200 kV, imaging modes: bright/dark field, STEM and magnification range 2k-5000kx. Original images had different resolutions, 4K and 2K images were randomly cropped to fit dimensions 512x512 px. Images that did not require dimension changes were left in their original size. The initial training dataset consists of 100 images and the appropriate number of binary masks. Images are made up of various elements and samples (Au, C, Zn, Bronz, KCl, Spinel, Az, Grid).

For image segmentation were created new larger training datasets. With the use of Augmentor [2] tool were created new larger datasets containing 1000, 5000, 10 000 images. This tool can use various image modification functions to generate the specific amount of output images. For testing purposes, three available functions were used – Rotation(R), Distortion(D) and Zoom(Z) [2]. There was also created a naming convention, to keep track of testing data. Each training dataset was named by an abbreviation of functions that were used for its generation and a numeric value that represents a number of images in the corresponding dataset.

- **R 1000, R 5000:** the probability of rotation (0.7), the range of image rotation (-25°, 25°), the rotation angle is relative to the centre of the image
- **D 1000, D 5000:** the probability of random distortion (0.9), distortion grid (8x8), the magnitude of distortion(2), elastic distortion is randomly performed in x and y-axis

- **Z 1000, Z 5000:** the probability of zooming (0.3), range of zooming (1.1, 1.6)
- **RDZ 1000**, **RDZ 5000**, **RDZ 10 000:** functions used precisely same parameters setups as in individual function usage above.

2.2 U-NET

The U-net has an architecture of a "fully convolutional network". Authors modify the design in such a way [1] that can work with very few training images and provide more precise segmentation. The successive layers enhanced the traditional contracting network design (repeated 3x3 convolutions), where polling (2x2 max pooling) was substituted by upsampling (2x2 up-convolution). For better localisation of high-resolution features, the information from contracting path is assembled with the upsampled output in a successive convolutional layer. Final layer (1x1 convolution) serves for mapping the component feature vector to the desired number of classes. The network consists of 23 layers. Keras implementation was chosen for testing purposes [6].

2.3 MODEL TRAINING

Pre-processing step consists of resizing all images into defined size, creating a binary mask (background and area of interest), enlarging training datasets with the Augmentor tool [2], converting it into *.npy array format. The training ran for ten iterations with the batch size of 1 in 1000 size datasets and 4 in rest. In every iteration, the model was computed from randomly chosen 80 % of test data and validated on the rest 20 %. As a model optimiser the Adam algorithm (learning rate = 0.0001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 1\times10^{-7}$, decay = 0) was used and as an optimisation score function - binary cross entropy loss function was used. Ten computational iterations spat out the best model. This model's task was to predict segmentation masks from data that network never saw before.

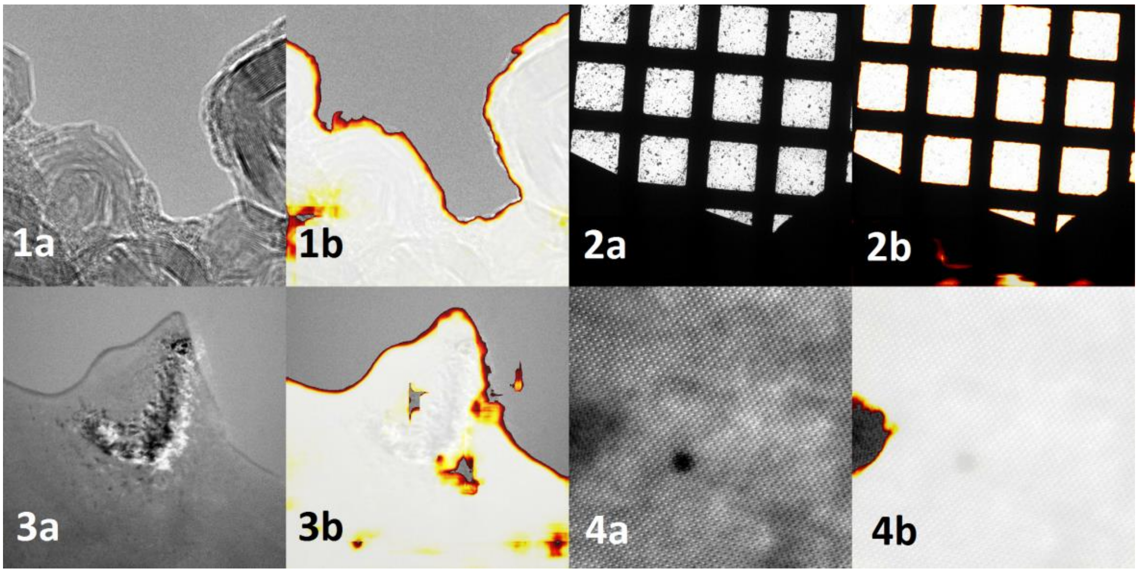

Figure 1: Segmentation on unknown data - visual results, **a** original images, **b** U-net output segmentation masks combined with the original images, transparency 80%, **1a, b** is graphitized carbon, **2a, b** is sample of cross-grating, **3a, b** is spinel sample, **4a, b** is silicon layer.

3 RESULTS

Based on the accuracy and loss function results from each model computation, the right approach was to select the RDZ 10 000 dataset for prediction purposes. With the inner model accuracy 0.99, all suggested that this model could achieve excellent results in segmentation tasks for unknown data. Figure 1 demonstrates visual results from four randomly chosen test images for segmentation task. Images with the same label number but different label letter create pairs. On b - label side is visualised grade to the class 1 - region of interest (white) and 0 - background(black). Mask 1b contains small FN segment in the third quadrant, 2b has FP artefacts in the lower part of third and fourth quadrants, 3b suffers from several small FN segments, mainly in the fourth quadrant and finally, the 4a mask was not able to detect small black burn hole that resulted in FP segment. The Dice similarity coefficient (DSC) [3] was used to evaluate the performance of segmentation, mean value 0.710 (range, 0.518-0.869), test set contained six samples.

4 DISCUSSION

This work successfully tested U-net neural network topology [6] for segmentation task on provided TEM image data. This specific neural network implementation was able to achieve high accuracy scores in segmentation task. Such high score values confirmed the original intention of authors of this topology [1] that even with the initially small amount of labelled data is possible to achieve excellent segmentation results. The RDZ 10 000 dataset was used for the primary segmentation model. The calculated DSC score from the test data indicates that the U-net neural network has the potential to achieve excellent segmentation results. From visual perspective is evident that the model is creating false positive and false negative artefacts. Manual segmentation of training data that was performed by two experts without the crosscheck validation could cause such artefacts. There are several possibilities how to avoid this problem. One possible solution of these artefacts is to add more variety in the original training dataset; the second solution could be to invite more experts and crosscheck their manual segmentation. The third solution would be to modify the network design [1]. The safest modification could be done in the final layer where different activation function can make a significant difference. Despite current results, this work has the potential, with right modifications, to achieve much higher DSC scores in segmentation tasks.

ACKNOWLEDGEMENT

I would like to express my thanks to several people from Thermo Fisher Scientific for providing me with image data, essential information and computational resources. Special thanks to Jan Pokorný and Ondřej Machek.

REFERENCES

[1] Ronneberger O., Fischer P., Brox T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham. Available at: https://arxiv.org/abs/1505.04597

[2] Marcus D. Bloice, Christof Stocker, and Andreas Holzinger. 2017. Augmentor: An Image Augmentation Library for Machine Learning, arXiv preprint arXiv:1708.04680, Available at: https://arxiv.org/abs/1708.04680

[3] Zou, Kelly H., Simon K. Warfield, Aditya Bharatha, et al. 2004. Statistical validation of image segmentation quality based on a spatial overlap index1. *Academic Radiology* [online]. 11(2), 178-189. Available at: https://tinyurl.com/y7me78tx

[4] Russakovsky, Olga, Jia Deng, Hao Su, et al. 2015. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision [online]. 115(3), 211-252. Available at: http://link.springer.com/10.1007/s11263-015-0816-y

[5] Long, Jonathan, Evan Shelhamer and Trevor Darrell. 2015. Fully convolutional networks for semantic segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. IEEE, 3431-3440. Available at: https://tinyurl.com/ycwbz4rf

[6] Mocko, Stefan, Unet, 2017. GitHub repository. Available at: https://tinyurl.com/y8ypdgz6